INT. J. REMOTE SENSING, 1994, VOL. 15, NO. 17, 3473-3491

# Characteristics of composited AVHRR data and problems in their classification

A. MOODY

Department of Geography, University of North Carolina, Chapel Hill, NC 27599-3220, U.S.A.

and A. H. STRAHLER

Center for Remote Sensing, Boston University, 725 Commonwealth Avenue, Boston MA, U.S.A.

**Abstract.** Unsupervised classification procedures were applied to a temporal sequence of fifteen bi-weekly composited NDVI images produced from AVHRR LAC data. Individual examination of the input images appeared to show substantial contamination due to clouds which persist through the compositing period. These apparent cloud features dominated the results of the clustering procedures. The composites also include large numbers of far off-nadir pixels. This causes severe spatial smoothing and produces a blurred image appearance. Further combining the data to monthly composites largely eliminated the cloud cover problem, but did not necessarily reduce the frequency of large view zenith angles. Preprocessing of high temporal frequency, low spatial resolution data such as that provided by AVHRR and the planned EOS MODIS instrument must more effectively remove the effects of clouds, correct for anisotropic scattering from the atmosphere and bi-directional reflectance from the surface, and should be biased towards the selection of near-nadir measurements. The design and processing procedures for MODIS will reduce problems associated with atmospheric effects and the geometric distortion of pixels, while enhancing the detection and screening of clouds.

## 1. Introduction

The availability of global satellite data on a frequent basis is a natural prerequisite to monitoring land cover and land-cover change on a global scale. Land-cover information is crucial for understanding processes such as global carbon cycling, terrestrial primary productivity, hydrologic cycles and terrestrial energy balance (Tucker *et al.* 1985). Timely information about vegetation cover is also important for monitoring rates of change of biotic resources, forecasting regional crop yields, and responding to famine brought on by drought conditions (Tucker and Choudhury 1987).

The only current system capable of providing this type of coverage is the Advanced Very High Resolution Radiometer (AVHRR) aboard the National Oceanic and Atmospheric Administration (NOAA) series of polar orbiting, meteorological satellites. AVHRR data have been the focus of abundant research on land cover characteristics and vegetation dynamics over the past decade. This research activity takes on new significance as Earth scientists prepare for the availability of data from the Moderate Resolution Imaging Spectrometer (MODIS) (Salomonson *et al.* 1989), which is presently scheduled for flight on the NASA Earth Observing

System (EOS) morning and afternoon satellite platforms. MODIS will be the primary EOS sensor for large-area land cover studies. While the MODIS instrument will be vastly superior to the AVHRR sensors for monitoring terrestrial processes, an understanding of the strengths and weaknesses encountered in AVHRR data have implications for the processing of MODIS data for studies of global land cover and land-cover change.

A great deal of research activity has been devoted to developing techniques for using AVHRR data for continental or regional-scale studies of land cover. Much of this work has focused on preprocessing techniques, such as the development of vegetation indices and image compositing procedures. Maximum value compositing of normalized difference vegetation index data (NDVI) (Townshend 1994) has proven useful for reducing cloud contamination and atmospheric effects (Tarpley *et al.* 1984, Holben 1986). Other work has focused on methods for exploiting the relatively high temporal resolution of AVHRR data to extract land-surface information based on the time varying properties of the surface cover (Townshend *et al.* 1987, Lloyd 1990, Loveland *et al.* 1991). Most studies which have employed AVHRR data to assess land-cover type have involved the compositing of NDVI data to reduce clouds and atmospheric effects.

This paper focuses on specific characteristics of a series of composited AVHRR-LAC (Local Area Coverage) NDVI images which were used in an unsupervised clustering of a region of the northeastern United States. We discuss some of the effects of clouds and off-nadir viewing on the clustering results and what implications they might have for the automated production of global land-cover datasets.

## 2. AVHRR data

### 2.1. *Basic issues*

The high temporal frequency of AVHRR data presents the possibility of near-continuous monitoring of land cover relative to the rate at which most surface-altering processes occur. This high temporal resolution represents the greatest strength of the system, as it allows the assessment of land-cover properties based on the spectral behaviour of the surface through time. However the development of multi-temporal AVHRR datasets which capitalize on this frequent sampling rate has proven difficult in part because of frequent cloud coverage over large areas of the Earth's land surface, and secondly because of geometric distortion and atmospheric contamination at large view zenith angles associated with the edges of the across-track scan. Production of high quality datasets from the AVHRR system represents, in part, a trade-off between the selection of near nadir-view images and selection of cloud-free images. The temporal frequency of a time series of images is then limited by the need to satisfy these quality constraints.

### 2.2. *View zenith angle*

The wide swath width of the AVHRR instrument, which allows high frequency temporal coverage, also results in degraded spatial resolution and geometric distortion of pixels at the edges of the scan. At the edge of the scan (i.e., $55^\circ$ off-nadir) the 1·1 km by 1·1 km nadir IFOV (instantaneous-field-of-view) is enlarged to 2·4 km by 6·5 km (Goward *et al.* 1991). Not only does the ground area per pixel increase at large view zenith angles, but pixel overlap is greatly increased and the IFOV rotates as the sensor looks off-nadir. These distortions confuse the interpretation of off-nadir scenes.

Moreover, Holben and Fraser (1984) noted that the afternoon NOAA satellite scans near the principal plane of the sun. Off-nadir measurements will be particularly affected by atmospheric scattering if they are sampled in the back-scatter direction, which tends to reduce the value of the NDVI because the atmospheric back-scatter effect is stronger in the visible than in the near-infrared wavelengths (Goward *et al.* 1991). Because of their proximity to the principal plane, off-nadir measurements from the afternoon instruments are also influenced by the bidirectional reflectance properties of the surface because of their proximity to the principal plane (Goward *et al.* 1991). The degree to which this influence occurs, and the effect that it will have on NDVI values is dependent on the surface characteristics.

The above considerations have implications for the production of AVHRR time-series datasets. When overlayed in a temporal sequence, a series of images, each composed of pixels sampled from a variety of view angles, may suffer misregistration and include nonlinear atmospheric and surface reflectance effects. Townshend *et al.* (1992) examined the effect of a variety of levels of misregistration on apparent changes in NDVI values between image pairs. Their results suggest that, for some areas, misregistration must be kept as low as 0·2 pixel to obtain 10 per cent errors or less from the true changes in NDVI.

### 2.3. *Clouds*

As mentioned above, the primary advantage of AVHRR data lies in their high temporal resolution. While daily coverage is theoretically available, many of these data are contaminated by clouds and cloud shadows, and different parts of the same scene may be cloud-covered on a series of consecutive days. The analysis of a series of single-date images for land-cover studies is, therefore, problematic as many contaminated pixels may be included in each image. Numerous techniques for separating clear radiances from cloud-contaminated values allow screening for clouds in single-date images. Most cloud detection algorithms have used the visible/near-infrared window and the longwave infrared window to obtain cloud information. Such algorithms include the Air Force Real-Time Nephanalysis Program (RTNEPH) and the International Satellite Cloud Climatology model (ISCCP) (Isaacs 1993). This approach is limited as clouds may be indistinguishable from bright and/or cold surfaces, thin cirrus clouds may be difficult to detect, and partly cloudy conditions are not reliably recognized (Isaacs 1993). Newer approaches include decision-tree algorithms which employ multispectral data including visible, near-infrared and longwave infrared bands. These types include the Air Force Tactical Nephanalysis Program (TACTNEPH) which uses multi-source data and dynamically adjusts to changes in data quality and availability, and the Clouds from AVHRR algorithm (CLAVR). Decision steps may include tests for thermal infrared thresholds, spatial cohesion of thermal data, reflective thresholds, ratios of bi-directional reflectance effects, and split window tests (Saunders and Kriebel 1988). These approaches are more successful at adjusting for problematic background conditions and allow the detection of optically thin cirrus clouds (Davis *et al.* 1988, Isaacs 1993). While there are many options for the detection of clouds in AVHRR data, cloud masking results in the production of single-date images with many missing values. This makes between-date comparison of pixel values difficult. For this reason, most assessment and monitoring of land-cover characteristics with AVHRR data has involved image compositing to produce time-series of imagery which have reduced temporal resolution, but more complete coverage of the surface.

## 3. Compositing

### 3.1. *Basic method*

The maximum-value compositing procedure is regularly used to produce time-series of AVHRR data which are relatively cloud free and have reduced atmospheric contamination. In this process, a sequence of registered images is compared and an output image is produced that consists of the greatest NDVI value observed among those in the image sequence for each pixel location. The compositing period is typically fixed, and each location receives a value from some date within that period. While this method produces certain desirable characteristics, it guarantees nothing about the quality of the resulting data except that it consists of the greatest NDVI values encountered within the compositing period.

Compositing has proven reasonably effective for selecting against cloud-contaminated measurements (Holben 1986). Large dense clouds will produce very low NDVI values, as they reflect nearly equally in both the red and NIR wave bands. Similarly, compositing of NDVI values can select against sub-pixel clouds and thin clouds. Small dense clouds will reduce the NDVI for a given pixel as they are averaged in with the underlying land cover to produce a digital count for that pixel. Thin clouds may also reduce the NDVI values over an area because they cause greater red reflectance, but still allow some of the signal from the scene below to pass to the sensor. These cloud types may not be readily observed in the reflected or thermal data at AVHRR resolutions, but may influence the NDVI values considerably (Goward *et al.* 1991).

### 3.2. *Directional effects*

As noted above, the NDVI can be influenced by the non-linear nature of atmospheric scattering and surface reflectance. That is, due to the anisotropic reflectance characteristics of the atmosphere and the surface, the value of the NDVI for a given ground location is dependent on the viewing and illumination geometry associated with a specific measurement. The relative importance of these two effects depends not only on observational geometry, but on the current atmospheric conditions and the specific bidirectional reflectance properties of the surface features. These effects may be further confounded by the increased atmospheric path length at off-nadir viewing positions. Two problems may then arise. Since NDVI values, produced from measurements sampled under different observational geometries, can be biased by the anisotropic reflectance characteristics, they may, in turn, generate errors in any surface information which they are used to generate. Alternatively, if measurements from multiple viewing and illumination conditions are compared within the maximum value compositing period, then the compositing procedure may preferentially select for off-nadir pixels. Off-nadir measurements are relatively undesirable as they exhibit geometric distortion, are affected by increased atmospheric pathlength, and their brightness values may be directionally biased due to atmospheric scattering and BRDF (bidirectional reflectance distribution function) effects.

Results from Holben and Fraser (1984) and Holben (1986) suggest that NDVI values are greatest near-nadir in the forward-scattering direction. This is because the NDVI is reduced in the back-scatter direction by atmospheric effects. Under an assumption of Lambertian surface reflectance, simulations of the NDVI have a convex shape along a scan line and, in the back-scatter direction, are greatest near the nadir position. A simulation which incorporated measured surface directional reflectances demonstrated a slight convexity of NDVI values at the surface along a scan line but suggest that, under most atmospheric conditions, the greatest NDVI value in the back-scatter direction is still not far from nadir (Holben 1986). However, results presented by Gutman (1991) indicate that the maximum value compositing procedure, when performed on a 40 km resolution dataset, is strongly biased toward measurements sampled from view angles in the forward-scattering direction. Due to the coarse resolution of the data used it is not clear what implications this has for the processing of AVHRR-LAC data. Similarly, Deering and Eck (1987) presented an analysis of surface-based measurements sampled in the principal plane during high solar zenith angles. Their results demonstrate a consistent elevation of NDVI values in the forward-scatter direction for three surface covers under both clear and hazy skies. An important exception was the increase in NDVI in the back-scatter direction for soybeans under clear-sky conditions. Because of the large solar zenith angles, and thus the strong internal shadowing within the vegetation cover, it is uncertain to what degree these results relate to the sampling of AVHRR data from afternoon platforms.

### 3.3. *Compositing interval*

The length of the compositing interval presents a trade-off between maintaining high temporal resolution and producing high quality images. Shorter compositing periods may allow improved resolution of surface cover dynamics, but the chance of acquiring cloud-free near-nadir measurements will be reduced. The idea, then, is to select a compositing period which is short enough to resolve the temporal dynamics of the ground cover, but long enough to allow the collection of high quality measurements.

The ideal time period for compositing NDVI imagery will depend upon the timing, frequency and duration of cloud cover over an area, the rate of change of local vegetation conditions, and the sensitivity of the vegetation to variable environmental conditions. For example, areas, such as an equatorial rainforest, might demand a very long compositing period perhaps, so long as to render the resulting image too discontinuous to be of value. In contrast, an area such as the Sahelian zone in northern Africa is so sensitive to environmental conditions and so variable from one year to the next, that a temporal sampling frequency of 7–10 days or less might be necessary to capture changes in land cover and adequately characterize the temporal dynamics of the vegetation.

### 3.4. *Registration*

The maximum-value compositing technique requires accurate registration among the images within the compositing period. The problem of registering a multi-temporal sequence is compounded because the images within each period are sampled from different view zenith angles. This leads to an inconsistent geometry of the AVHRR IFOV among the images that need to be co-registered. Aside from the effect on multi-image registration, the enlargement, overlap and rotation of the IFOV causes a spatially blurred appearance of off-nadir images. Both geometric distortion due to off-nadir viewing and misregistration of input data can produce composited images that have a spatially blurred or smeared appearance. If this blurring is due to registration error, then subsequent use of the composited images to perform land-cover characterization could lead to misclassification, as pixels from different locations on the ground would be compared to each other in the compositing procedure. If such misregistration occurred within compositing periods, then there would also be error between periods. This would further compromise the classifier performance because pixel time-trajectories would not necessarily track consistent ground locations. If observed blurring of composited images is only due to IFOV geometry, then this amounts to a spatial smoothing of the data. The effect this would have on the output of a classification procedure would depend on the spatial pattern of the land-surface features in the scene. While blurring of the composited images is frequently referred to in the following sections, no attempt is made to separate between that which results from IFOV geometry and that which results from registration error.

## 4. Study

### 4.1. *Objectives*

We examined a temporal sequence of nineteen biweekly composite NDVI images produced from AVHRR-LAC data over a period extending from March to December 1990. These data are available on CD-ROM for the conterminous U.S. in the AVHRR bi-weekly composite dataset produced by the U.S. Department of Interior, Geological Survey at the National Mapping Division, EROS DATA Center. This dataset was produced using measurements from the AVHRR instrument aboard NOAA-11. This platform is in an ascending, near-polar orbit with a 14:30 equatorial crossing time. The bi-weekly composites were produced using the maximum value compositing procedure. The region of interest covered a 512 pixel by 512 pixel image segment and included most of New England.

Our intent was to explore various methods for characterizing land cover using AVHRR composited-NDVI data. Techniques included unsupervised clustering using a subset of the sequence of nineteen images (e.g., Loveland *et al.* 1991), clustering of principal component images, and clustering based on derived variables such as the length of the growing season, maximum NDVI value, date of maximum NDVI, and date of the onset of greenness (e.g., Lloyd 1990). During the processing stage, it became apparent that the dataset was affected by cloud cover, which strongly influenced the classification results. A closer examination of the data also revealed a severe loss of detail, or spatial blurring, in many of the composited images associated with the inclusion of large numbers of off-nadir pixels through the compositing procedure. This paper discusses how clouds appeared to affect the clustering results, the degree to which off-nadir pixels are incorporated into the composited images, the potential consequences of this inclusion, and the effect of further compositing on cloud-covered and off-nadir pixels.

### 4.2. *Cloud effects*

Several distinct features appeared in all clustered images which were, at first, attributed to pronounced terrain features, including the Adirondack Mountains, the Hudson and Connecticut River Valleys, the White Mountains and the Green Mountains. Upon closer examination of the original input images, it was determined that these features in the classified output were closely related to a pattern of clouds in a single composited image. Figure 1 (*a*) is an image of the sixth bi-weekly composite of the study area, derived from images acquired from 11–24 May 1990. This composite is clearly affected by clouds, which are especially apparent as dark patterns in the regions of the Adirondacks and the White and Green Mountains. Figure 1 (*b*) is the result of an unsupervised classification using a clustering

(a)

(b)

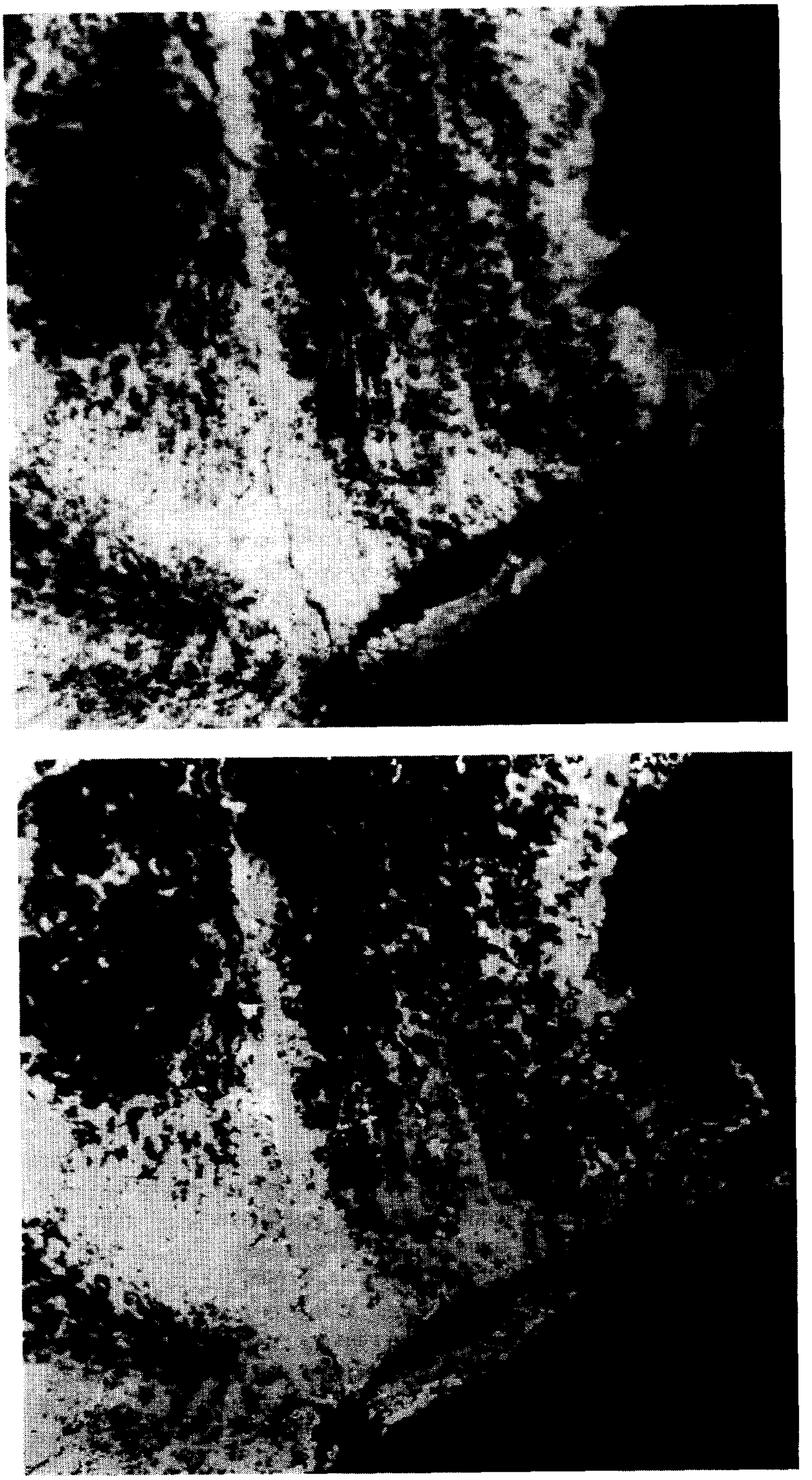


Figure 1. Comparison between (a) NDVI composite #6 (representing 11–24 May 1990), and (b) results from an unsupervised clustering using 15 of the 19 available bi-weekly composited images. Composite #6 strongly influenced the final clustering result, as is evident from the similarity of the spatial patterns in the two images.

algorithm applied to images 3 through 17 of the 19 image sequence, which includes the composite of 11–24 May. A close examination of these two images reveals that not only are the main features very similar, but finer details are also alike. In this case, the clustering algorithm, which is of the ISOCLAS type (Kan 1972), differentiated many clusters based on the variations in NDVI values of the sixth composite, which were influenced by cloud cover.

### 4.3. *View zenith angle effects*

The selection of a large number of off-nadir pixels produces a badly smeared image appearance which obscures natural boundaries and results in an image with reduced spatial variability. About half of the images in the bi-weekly sequence, we examined, exhibit noticeable, and sometimes severe smoothing of this nature. Images created to display the view zenith angle of each pixel in the composites demonstrate that large numbers of off-nadir pixels are routinely included by the maximum value compositing procedure. This is also shown in histogram displays of viewing angle frequency. A pair of images of differing quality were selected for comparison through examination of the NDVI composite images and the associated view angle images and histograms. These are displayed in figure 2. The satellite zenith angles presented here are calculated relative to the surface normal and, therefore, extend beyond the 55° consistent with the AVHRR scan angle.

An examination of the image pairs in figure 2 demonstrates the relationship between off-nadir viewing and image blurring. Figures 2(*a*) and (*b*) are the composited NDVI images numbers 2 (16–29 March) and 4 (13–26 April), respectively. Visual comparison of this image pair shows considerably reduced sharpness of detail in composite number 2. This is especially apparent in the relative clarity of the northern coastline and drainage features between the two images where composite number 2 appears smoothed relative to number 4. Notice in particular the entire northeastern portion of the image. Portions of composite 2 may also be contaminated by clouds. Figures 2(*c*) and (*d*) are images representing the satellite zenith angles associated with composite images numbers 2 and 4, respectively. In these images, the darker tones represent farther off-nadir, eastward-looking viewing angles. It is apparent that large areas of composite number 2 were sampled at large view zenith angles in the eastward direction (dark areas of figure 2(*c*)). By comparison, figure 2(*d*) is composed of predominantly intermediate tones (indicating near-nadir sampling) and does not demonstrate the large patches of off-nadir angles as seen in figure 2(*c*). Notice that the worst loss of detail in figure 2(*a*) is associated with the large areas of figure 2(*c*) that exhibit extremely eastward looking view zenith angles. The satellite view angle histograms associated with composites 2 and 4 are shown in figures 2(*e*) and (*f*) respectively. Figure 2(*e*) shows the predominance of extremely eastward looking view angles in composite number 2. By comparison, figure 2(*f*) indicates that most of the values in composite number 4 were selected from relatively near the nadir viewing position. Notice the large mass of values at the extreme eastward-looking angles associated with the water pixels in composite number 4. Figure 3(*a*) (composite number 8) is of interest as it represents the view zenith angles of an image for which the land values were derived almost entirely from one overpass in the compositing period.

### 4.4. *View zenith angle bias*

For this dataset, the direction and degree of off-nadir viewing is influenced by the location of the receiving stations relative to the location of interest on the ground.

Figure 2 (*a*)

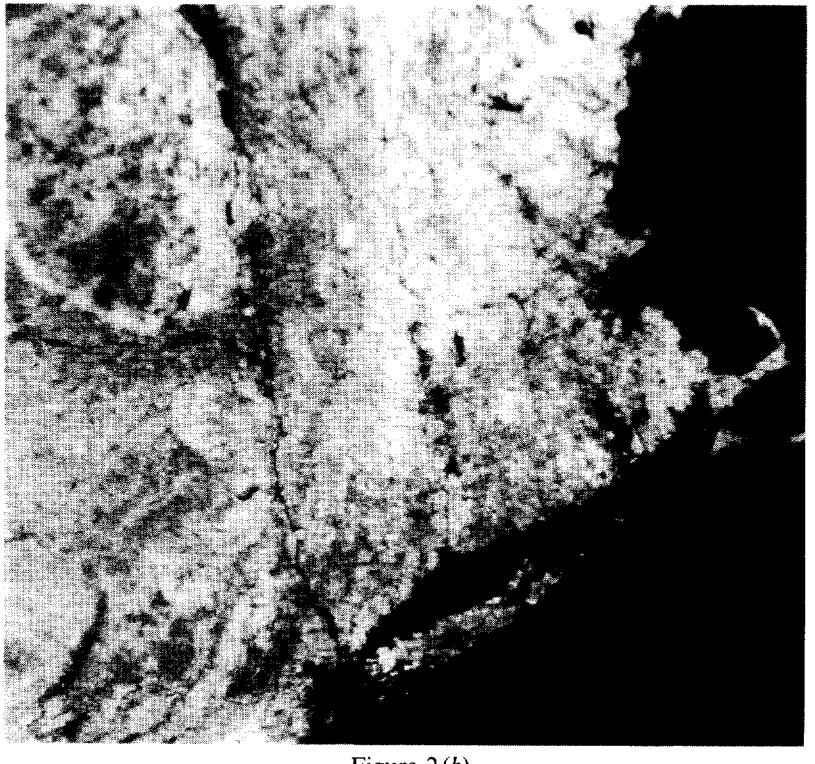


Figure 2 (*b*)

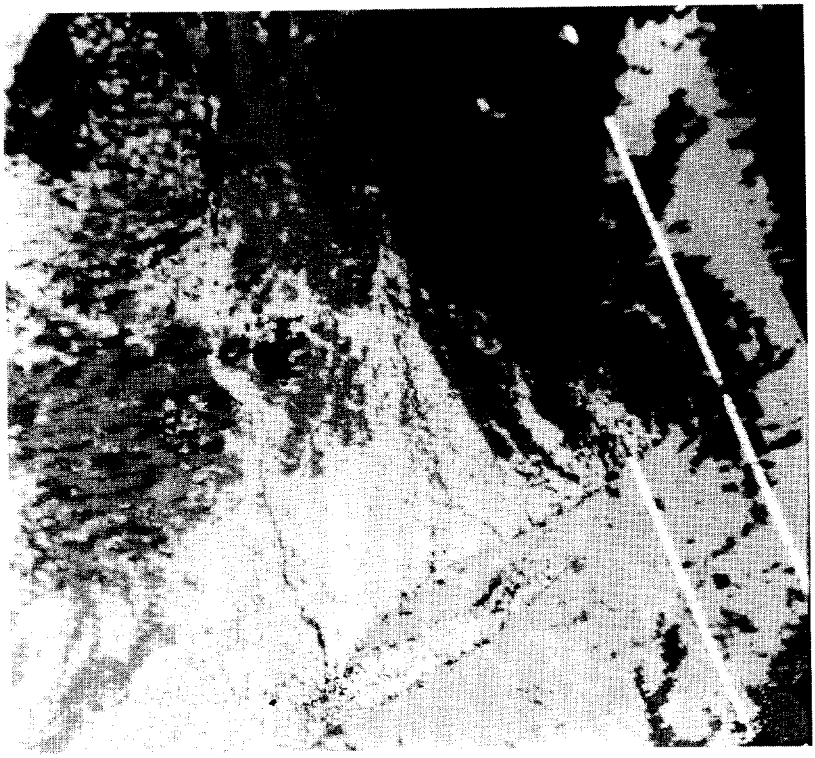
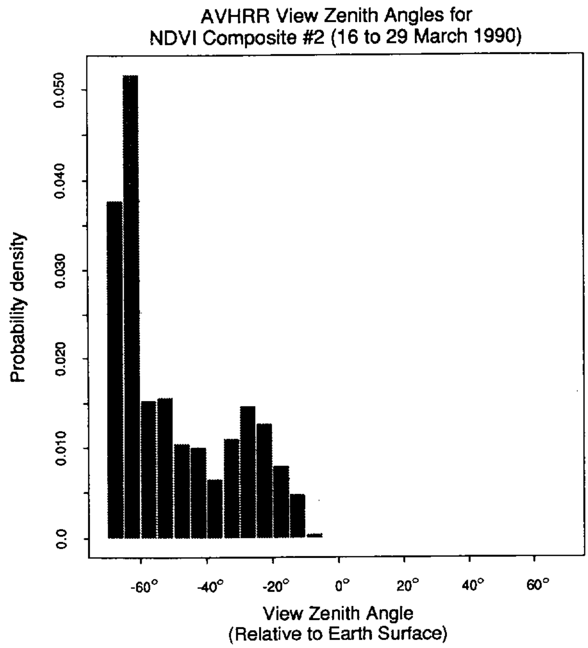
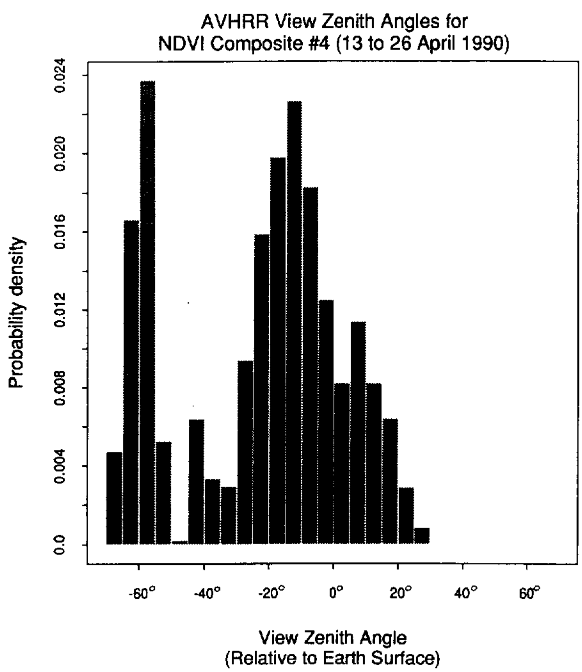


Figure 2 (c)

Figure 2 (d)




(e)




(f)

Figure 2. Comparison between NDVI images, view angle images, and view angle distributions for NDVI composite #2 (16–29 March 1990); and #4 (13–26 April 1990). Figures 2 (a), (c) and (e) represent the NDVI image, the view angle image and the view angle histogram for composite #2, respectively. Figures 2 (b), (d) and (f) similarly correspond to composite #4. View angles are represented relative to the surface normal.

This effect can be observed by comparing viewing angle images of the east coast and the west coast of the United States. West coast images are mostly comprised of values acquired at a westward look angle, and viewing angle values approach nadir as pixels are examined closer to the eastern edge of the image. East coast images display the opposite pattern; that is, they are acquired largely at an eastward look angle and viewing angle values approach nadir at pixel locations closer to the western edge of the image. This selection is presumably due to the location of the data receiving station at Sioux Falls, South Dakota, roughly midway between the two coasts, and because of the coverage circle of the station's reception capability. That is, the tracking capability of the antenna is limited over the Atlantic, so it picks up the next orbit which is a nadir or eastward look at the east coast. A similar situation occurs over the west coast (Eidenshink, personal communication). This effect can be seen in figures 3 (*a*) and (*b*) which are the images of viewing angle values for locations on the east and west coasts of the United States, respectively. In figure 3 (*a*) (east coast), the brightest values are near nadir and the view angles become increasingly eastward-looking as the grey values darken. The nearest to nadir view is at the lower left of the image and the upper right is the farthest off-nadir in the eastward direction. Figure 3 (*b*) (west coast, centred on the San Francisco Bay area) can be interpreted similarly. Light values are near-nadir and the view angle becomes increasingly westward looking as the grey values get darker. Here, the nearest to nadir is at the right of the image and the values at the left are the farthest off-nadir in the westward direction.

### 4.5. *Recompositing*

Further compositing this dataset into monthly maximum NDVI images produced a dataset that appeared to be relatively cloud-free. The monthly composites were created by selecting the greatest NDVI value for each pixel in each adjacent pair of bi-weekly images. While this produced the brightest possible output image and, therefore, one that was relatively unaffected by clouds, it did not necessarily produce the best image in terms of image sharpness, or near-nadir sampling. If the brighter image of the two was comprised of many off-nadir pixels, then these pixels were selected and a blurry, off-nadir image was produced where a sharper image might have been possible. For example, a comparison of monthly composite number one (figure 4 (*c*)) with the two bi-weekly images from which it was derived demonstrates that the monthly composite looks much like the second image in the bi-weekly sequence. While the second image (figure 4 (*b*)) has greater NDVI values than the first (figure 4 (*a*)), the first image is clearly superior in sharpness of detail. These images are independently stretched, so direct comparison of brightness values cannot be made.

In general, the compositing procedure may select for off-nadir pixels under several circumstances. If, within the compositing period, cloud cover occurs over an area during times of nadir viewing but there have been non-cloudy off-nadir viewing opportunities, then the off-nadir measurements will be selected because their NDVI values have not been reduced by clouds. Second, if, during a green-up phase, the latest measurements in the compositing period are sampled from off-nadir viewing angles they may produce brighter NDVI values and, therefore be preferentially selected. Conversely, during senescence, off-nadir measurements will be preferred if they are the earliest samples within the period of compositing. Finally, the directional effects associated with the surface and the atmosphere may lead to

(a)

(b)

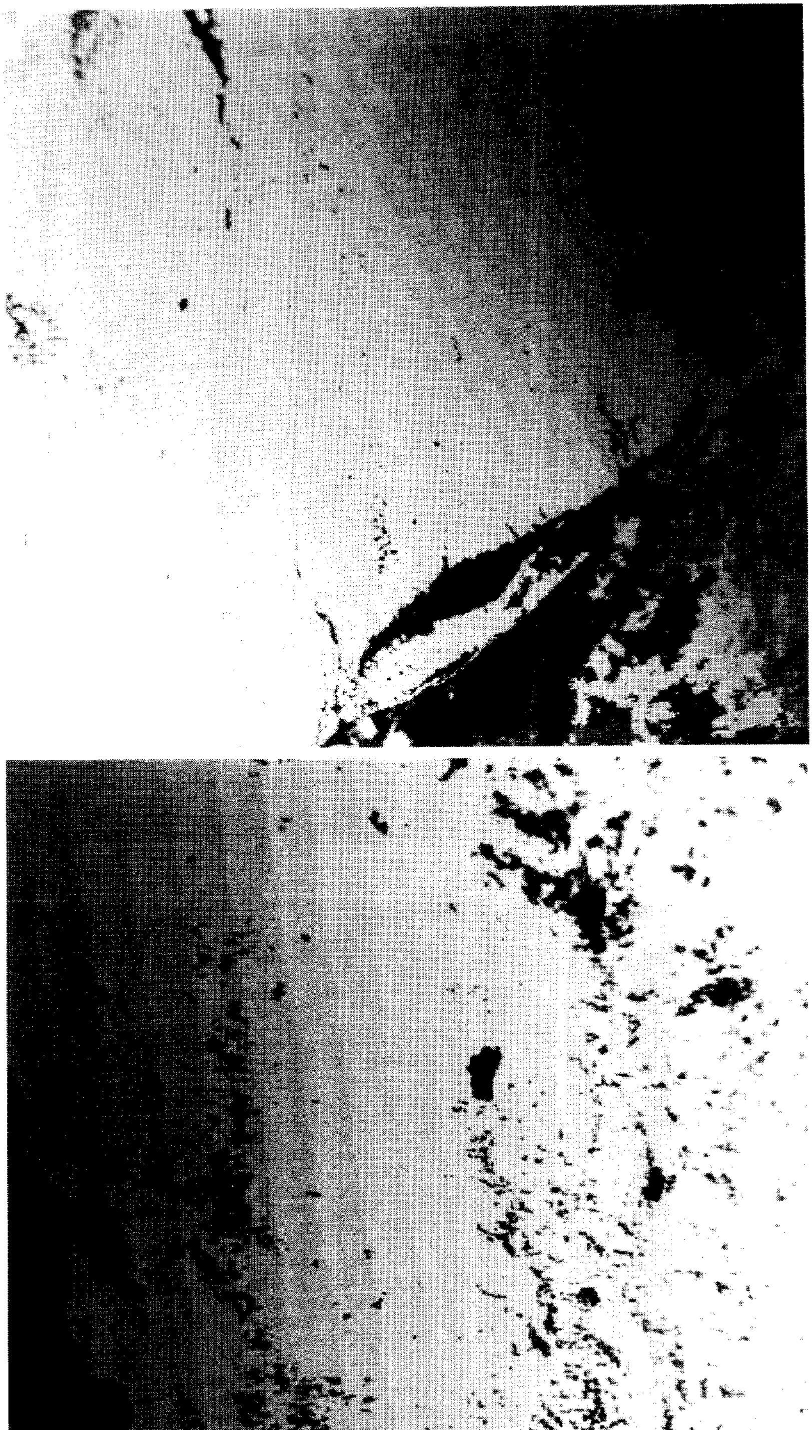
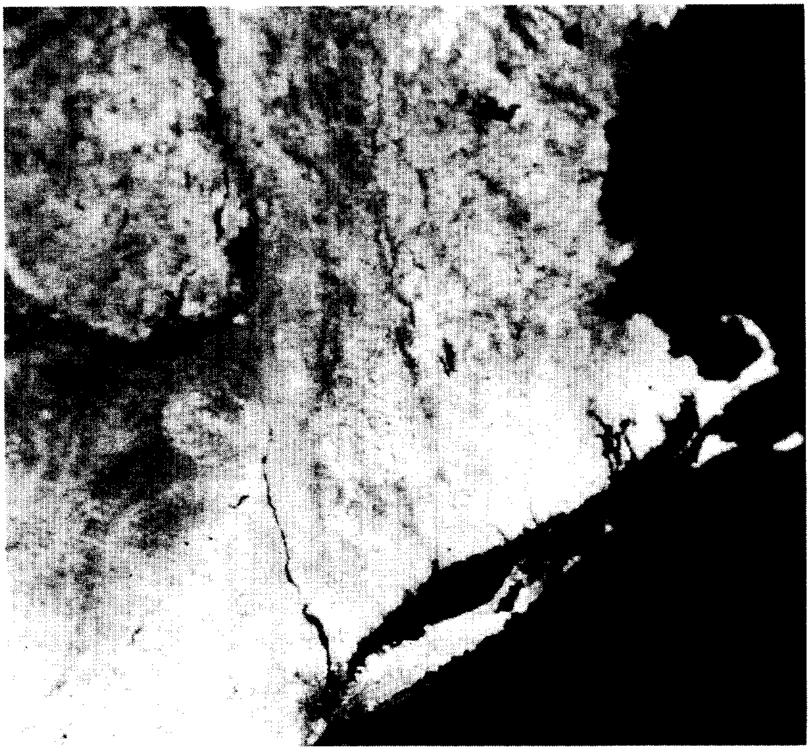
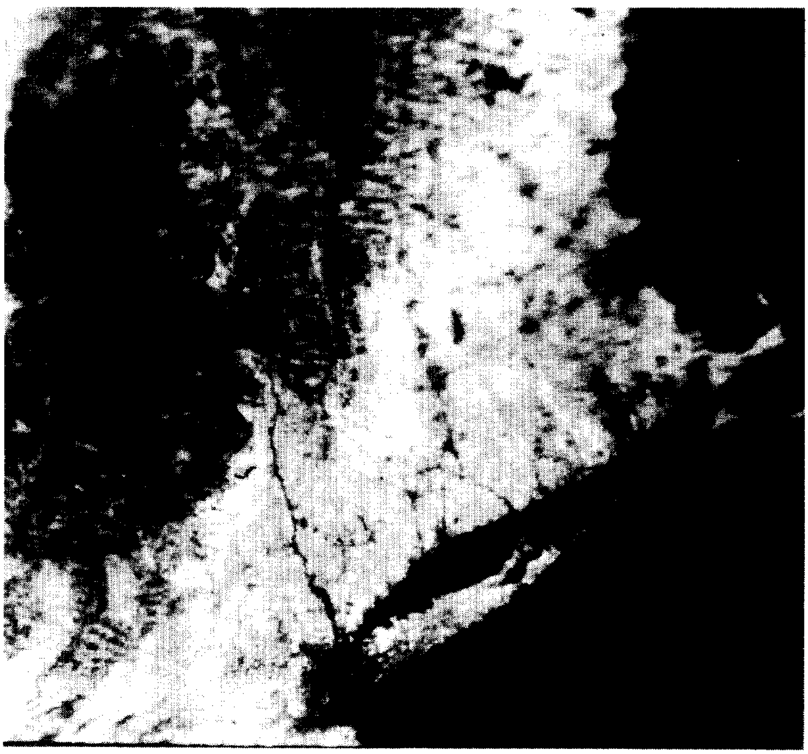


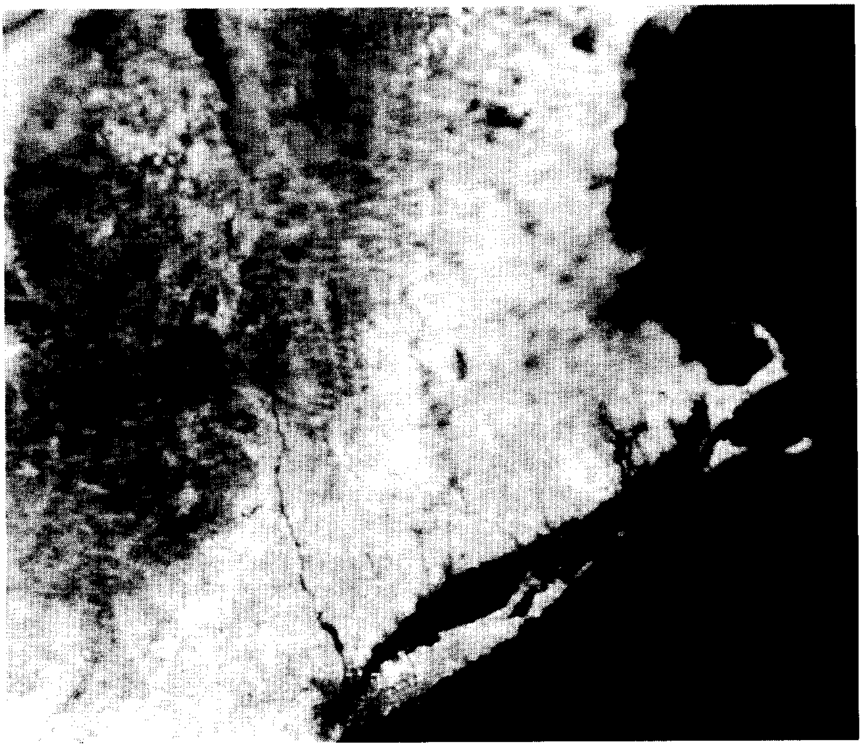
Figure 3. Demonstration of view angle bias using (*a*) the view angle image for composite #8 (8–21 June 1990), and (*b*) a west coast view angle image for composite #19 (7–20 December 1990) around the San Francisco Bay area. In (*a*) the brightest values represent the near-nadir view while the darker values are increasingly eastward looking. In (*b*) the darkest values are westward looking.



(a)



(b)



(c)

Figure 4. (a) NDVI composite #1 (2–15 March 1990); (b) NDVI composite #2 (16–29 March 1990); and (c) a combination of images #1 and #2 into a monthly composite.

greater NDVI values when sampled at large viewing angles, especially in the forward-scattering direction. These directional effects may compound or counteract the effects mentioned above.

It is assumed for two reasons that the data examined in this study were only minimally influenced by directional atmospheric and surface reflectance effects. First, as noted above, the location of the receiving station leads to a strong bias toward nadir or eastward-looking view zenith angles (back-scatter direction) for the east coast location. Examination of view zenith angle histograms for all 19 dates in the composited image sequence show, without exception, the greatest frequency of values near the nadir position or toward angles in the back-scatter direction. Thus, forward scattering measurements could not be preferentially selected as they are not generally present. Second, for the New England data, the instrument was not scanning close enough to the principal plane to sample the dominant components of anisotropic near-hotspot scattering from the atmosphere and bi-directional reflectance from the surface (Barnsley, personal communication).

## 5. Discussion

About half of the composited images for the region examined in this study appeared to suffer from at least partial cloud contamination. This indicates that the biweekly compositing period is too short to remove clouds effectively from this data. A longer compositing interval such as tri-weekly or monthly would be more appropriate. A longer period, however, may not allow the resolution of temporal land cover changes in more dynamic, rainfed agricultural areas. This dilemma limits the ability to use temporal dynamics as an aid in land-cover classification in areas where cloud cover is persistent and/or where critical temporal changes occur too quickly to be resolved between compositing periods. These issues suggest that a flexible compositing period which can be tailored to regional variations in land cover dynamics and meso-scale climate variability would better suit the purpose of monitoring land cover at the continental or global scale in an automated fashion. For example, a 5–10-day compositing period during early stages of processing could be recomposited as needed on a regional basis.

The fact that the cloud features in this dataset were readily apparent in the classified images suggests problems with interpreting the output of classification procedures using data composited over a short duration. This is especially true if automated classification procedures are employed which do not involve iterative tuning of algorithms or manual interaction with the data. For frequent production of global datasets of land cover and land-cover change, such automated processes will be necessary. It is likely that more reliable cloud-screening procedures which effectively eliminate cloud-covered pixels from composited images will be developed. This purpose might best be served by decision-tree types of cloud detection algorithms such as CLAVR (Davis *et al.* 1993) or that proposed by Saunders and Kriebel (1988). If compositing is used, cloud masks could be applied before the compositing procedure to eliminate cloudy pixels that might persist through the processing period.

The frequent selection of off-nadir pixels in the compositing procedure has implications for automated processing where, ideally, the highest quality sequences of data can be incorporated without close human interaction. The data used here include large numbers of far off-nadir pixels, which results in the loss of spatial detail and a blurry appearance over large areas of the composited images. This loss of detail may compromise the accuracy of classification output produced using these data. Further combining the data to monthly composites appears to reduce the inclusion of cloud dominated measurements, but off-nadir measurements will not necessarily be removed. This condition naturally constrains the ability to observe surface features such as lakes, drainage features and shorelines, and will introduce classification error at class boundaries.

The poor locational accuracy and varying ground resolution associated with off-nadir sampling make the task of resampling and registering images between dates complex and error-prone procedures. The quality of between-date registration currently possible with AVHRR composited data severely limits the accuracy with which land-cover characterization can be accomplished (Townshend *et al.* 1992). When producing composite images, examining time series of composites, or comparing images between years for the purpose of change characterization, there is a possibility that registered pixels may not actually represent the same surface location or the same ground area.

In order to acquire data with the temporal frequency necessary to produce global datasets, the off-nadir measurements resulting from the wide swath-width of systems such as AVHRR and the future MODIS sensor must be used. All other factors being equal, however, near-nadir views are clearly preferred for applications such as land-cover classification where resolution of spatial pattern and consistent pixel geometry are desirable characteristics. Preprocessing algorithms should, therefore, be designed to select for near-nadir measurements which also meet other quality constraints. The correction of BRDF and atmospheric effects is also a critical step for removing directional reflectance biases of the surface and atmosphere.

While the maximum value compositing procedure as it is typically employed has proven useful for reducing the influence of clouds, it is possible that improved techniques can be designed to produce a better balance between the selection of cloud-free and near-nadir pixels. For example, a filtering operation based on texture could be designed to screen against off-nadir pixels whose smoothed spatial properties would produce low spatial variance. In the compositing context, maximum NDVI and maximum variance procedures might be combined in a way which would preferentially select against clouds as well as off-nadir measurements. Another solution is to simply limit input pixels to those measured from some middle portion of the across-track scan. However, this would not only reduce the temporal sampling frequency of usable measurements, but would reduce sampling frequency the most for locations between the tropics. These tend to be locations where frequent measurements are most needed to produce cloud-free images. Reference to Goward *et al.* (1991, figure 7) shows that limiting view angles to $\pm 26^{\circ}$ would permit only about 10 samples per month at equatorial locations. Clearly, if the acceptable view angle range is limited, consistent global monitoring will be compromised.

One approach to resolve some of these problems might involve compositing AVHRR data on a per-pixel basis. Pixels would be individually screened for clouds and viewing angle effects and would need to meet pre-established minimum criteria in order to be selected. This would occur within some processing period, as in the maximum value compositing procedure, so that temporal aliasing would be held within periods and so that interannual comparisons would still be meaningful. This would allow the preferential selection of cloud-free pixels as well as near-nadir measurements, and would not necessarily equate the most reliable, or best, measurements with those that produce the greatest NDVI values.

This type of compositing will become practical with the advent of MODIS, planned for launch in 1998. In the context of NASA's role in supporting the global science community through EOS, MODIS is planned as the primary sensor for monitoring terrestrial and oceanic processes on a global scale. MODIS continues the heritage of high temporal frequency, low spatial resolution sensors such as the NOAA-AVHRR series. However, the superior design of the MODIS instrument and its processing system will alleviate many of the problems encountered in processing AVHRR data for land-cover applications.

The potential for thermal detection of clouds will be enhanced by the large number of thermal bands, as will cloud/snow discrimination by the shortwave infrared band at $1{\cdot}64\,\mu$m. Two 250 m bands in red and near-infrared channels will facilitate detection of clouds that are too small to appear at 500 or 1000 m resolution. Also, atmospheric correction of land imaging channels is planned, again enhanced by the MODIS band complement (King *et al.* 1992). Further, vegetation indices that are insensitive to atmospheric effects such as the Atmospherically-Resistant Vegetation Index (Kaufman and Tanré 1992) are being developed that will serve as alternatives to the NDVI. With proper correction and diminished atmospheric effects, the primary function of compositing is then simply to fill in cloud-covered areas and record the peak vegetation index value during the compositing period.

While off-nadir MODIS measurements will suffer from geometric distortion, they will not overlap or rotate as successive AVHRR IFOVs do. This, combined with the improved pointing knowledge of the EOS-AM and -PM platforms, and possibly optical ground navigation provided by the MISR instrument (Multiangle Imaging Spectral Radiometer, Diner *et al.* 1987), will allow for improved geo-referencing and image registration.

## 6. Conclusion

This paper represents an exploration into some of the issues whch may plague high volume, automated processing of regional- or global-scale data for land cover and land-cover change characterization. The problems encountered dictate a high degree of user interaction when processing data such as that provided by AVHRR. If temporal dynamics are to be used as a cornerstone of, or input to, large area land-cover characterization, it will be necessary to produce cloud-free, multi-temporal datasets with sampling rates which are appropriate at the regional level and which are corrected for the relationship between view geometry and reflectance measurements and pixel geometry. If compositing is used, procedures must be advanced to avoid the inclusion of off-nadir pixels and should incorporate flexibility of the compositing period based on image quality and/or the frequencies of scene dynamics.

## Acknowledgments

The authors gratefully acknowledge the reviewers for their helpful comments and suggestions. We also thank M. Barnsley of University College, London for calculating BRDF sampling geometry for the New England data. This work was supported by the National Aeronautics and Space Administration under contract NAS5-31369.

## References

Davis, P. A., Stowe, L. L., and McClain, E. P., 1993, Development of a cloud layer detection algorithm for the Clouds from AVHRR (CLAVR) Phase II code. *Proceedings of SPIE Conference on Passive Infrared Remote Sensing of Cloud and the Atmosphere, Orlando, Florida*; in press.

Deering, D. W., and Eck, T. F., 1987, Atmospheric optical depth effects on angular anisotropy of plant canopy reflectance. *International Journal of Remote Sensing*, **8**, 893-916.

Diner, D. J., Martonchik, J. V., Bruegge, C. J., Danielson, E. D., and Kimes, D. S., 1987, Remote sensing of aerosol and surface properties with a multidirectional imaging spectroradiometer. *Proceedings of IGARSS '87 Symposium, Ann Arbor, Michigan* (Ann Arbor:IGARSS), pp. 667-

Goward, S. N., Markham, B., Dye, D. G., Dulaney, W., and Yang, J., 1991, Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sensing of Environment*, **35**, 257-277.

Gutman, G. G., 1991, Vegetation indices from AVHRR: An update and future prospects. *Remote Sensing of Environment*, **35**, 121-136.

Holben, B. N., 1986, Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, **7**, 1417-1434.

Holben, B. N., and Fraser, R. S., 1984, Red and near-infrared sensor response to off-nadir viewing. *International Journal of Remote Sensing*, **5**, 145-160.

Isaacs, R. G., 1993, Remote sensing of cloud for defense and climate studies: an over-view. *Proceedings of SPIE Conference on Passive Infrared Remote Sensing of Cloud and the Atmosphere*, Orlando, Florida; in press.

Kan, E. P. F., 1972, The latest version of ISODATA(A)/ISOCLS. Lockheed Electronics Company, Inc., HASD, Houston, Texas, Technical Memorandum TM642-570.

Kaufman, Y. J., and Tanré, D., 1992, Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **30**, 261-270.

KING, M. D., KAUFMAN, Y. J., MENZEL, W. P., and TANRÉ, D., 1992, Remote sensing of cloud, aerosol and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **30,** 2-27.

LLOYD, D., 1990, A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing*, **11,** 2269-2279.

LOVELAND, T. R., MERCHANT, J. W., OHLEN, D. O., and BROWN, J. F., 1991, Development of a Land-Cover Characteristics Database for the Conterminous U.S. *Photogrammetric Engineering and Remote Sensing*, **57,** 1453-1463.

RUNNING, S., JUSTICE, C., HALL, D., HUETE, A., KAUFMAN, Y., MULLER, J., STRAHLER, A., VANDERBILT, V., WAN, Z., TEILLET, P., and CARNEGGIE, D., 1994, Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *International Journal of Remote Sensing*, **15,** 3587-3620.

SALOMONSON, V. V., BARNES, W. L., MAYMON, P. W., MONTGOMERY, H. E., and ONSTROW, H., 1989, MODIS: Advanced facility instrument for studies of the Earth system. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **27,** 145-153.

SAUNDERS, R. W., and KRIEBEL, K. T., 1988, An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, **9,** 123-150.

TARPLEY, J. P., SCHNEIDER, S. R., and MONEY, R. L., 1984, Global vegetation indices from NOAA-7 meteorological satellite. *Journal of Climate and Applied Meteorology*, **23,** 491-494.

TOWNSHEND, J. R. G., 1994, Global data sets for land applications from the Advanced Very High Resolution Radiometer: an introduction. *International Journal of Remote Sensing* **15,** 3319-3332.

TOWNSHEND, J. R. G., JUSTICE, C. O., and KALB, V., 1987, Characterization and classification of South American land-cover types using satellite data. *International Journal of Remote Sensing*, **8,** 1189-1207.

TOWNSHEND, J. R. G., JUSTICE, C. O., GURNEY, C., and MCMANUS, J., 1992, The impact of misregistration on the detection of changes in land cover. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **30,** 1054-1060.

TUCKER, C. J., and CHOUDHURY, B. J., 1987, Satellite remote sensing of drought conditions. *Remote Sensing of Environment*, **23,** 243-251.

TUCKER, C. J., TOWNSHEND, J. R. G., and GOFF, T. E., 1985, African land cover classification using satellite data. *Science*, **227,** 369-375.